AD-A264 306

An Algorithm for Simple and Complex Feature Detection:

from Retina to Primary Visual Cortex
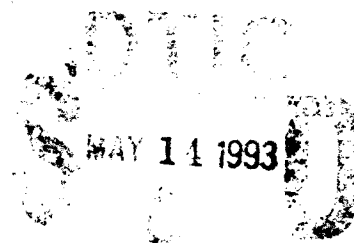
M. R. Blackburn

MAY 1 1 1993

93 5 1 090

93-10796

NRaD

Technical Document 2456
February 1993

An Algorithm for Simple and Complex Feature Detection:

from Retina to Primary Visual Cortex

M. R. Blackburn

NAVAL COMMAND, CONTROL AND
OCEAN SURVEILLANCE CENTER
RDT&E DIVISION
San Diego, California 92152-5000

J. D. FONTANA, CAPT, USN
Commanding Officer

R. T. SHEARER
Executive Director

PK

BACKGROUND

In the present paper, a hierarchy of functions is described for the processing of visual motion and pattern information culminating in descriptive sets of features. Processing elements that model simple feature detectors respond maximally only when a stimulus grating pattern with a specific orientation and spatial period is located in phase with the center and surround of the element's receptive field. Processing elements that model complex feature detectors respond maximally only when a stimulus grating pattern with a specific orientation and spatial period is moved in a specific direction through the element's receptive field.

The process of perception in biological visual systems involves both pattern analysis and pattern synthesis. Pattern analysis allows the discovery of distinguishing sets of features of input patterns that are useful for the discrimination of likely targets required for some specific behavioral response. Complex sets of features may be required to discriminate targets such as faces, while trigger stimuli that release sequences of species specific behaviors can be very simple features such as size, color, or motion (Ewert, 1987). Pattern synthesis allows prediction and selective filtering of targets from a complex and dynamically changing environment.

The division of perceptual processes into analysis and synthesis is somewhat artificial and done so for the convenience of modeling and discussion. This paper focuses on the processes of motion and pattern analysis that could provide the basic environmental information to a perceiving system. For an introduction to a complementary approach to the problem of pattern synthesis and prediction, see Blackburn and Nguyen (1989) and Blackburn (1990).

The neurobiological literature clearly indicates that feature detection, selective filtering and analysis are accomplished by using functions from multiple levels of the visual system. In mammals, much preprocessing of the visual information occurs in the retina and thalamus before it reaches the visual cortex where synthesis and recognition take place. It could be useful, in terms of neural modeling and the simulation of advanced perceptual processes, to provide as much biological fidelity as feasible of the early pattern analysis in the retina, thalamus and primary visual cortex to help define the more advanced perceptual processing algorithms. It is, therefore, the purpose of this research to examine methods of feature detection involving models of the motion and pattern analysis processes of the biological retina, thalamus, and primary visual cortex.

An additional neural structure, the superior colliculus (SC) of the brain stem tegmentum, also significantly determines the information available to the visual cortex. The SC, though not interposed in the major transmission pathway from the retina to the cortex, determines the visual input to the cortex through its control of saccadic eye movements. In the simulations presented below, all activity patterns are presented from computations that have occurred in the inter saccade intervals.

A schematic of the components of the mammalian visual system considered in the present

1

paper is given in figure 1. The components, except for the SC, will be discussed in the following sections. For a description of the model of the SC and its role in the system, see Blackburn (1993b).

ALGORITHM

CENTER SURROUND RECEPTIVE FIELDS

Two major streams of visual information flow from the retina to the cortex. These focus on the analysis of motion and pattern information respectively (Van Essen and Maunsell, 1983). Motion information is thought to be initiated in the fast conducting transient Y-type retinal ganglion output cells, while pattern information is suggested to be initiated in the slower conducting, sustained X-type retinal ganglion cells (Stone et al., 1979). Pattern analysis begins in the outer plexiform layer of the biological retina (Dowling, 1987). In that layer, center surround receptive fields are processed to provide a mechanism of contrast enhancement. We have explicitly modeled this mechanism by computing the local spatial derivative of the contrast gradient in the visual input (Blackburn, 1993a). Both on-center/off-surround and off-center/on-surround receptive fields are organized, providing effects that have been described by the now familiar two-dimensional difference-of-Gaussians (Rodieck, 1965). Currently, the activities of the bipolar elements[1] are passed directly to ganglion output elements. These exist as complementary pairs ($RX0_{i,j}$ and $RX1_{i,j}$)[2] covering the same locations in the visual field.
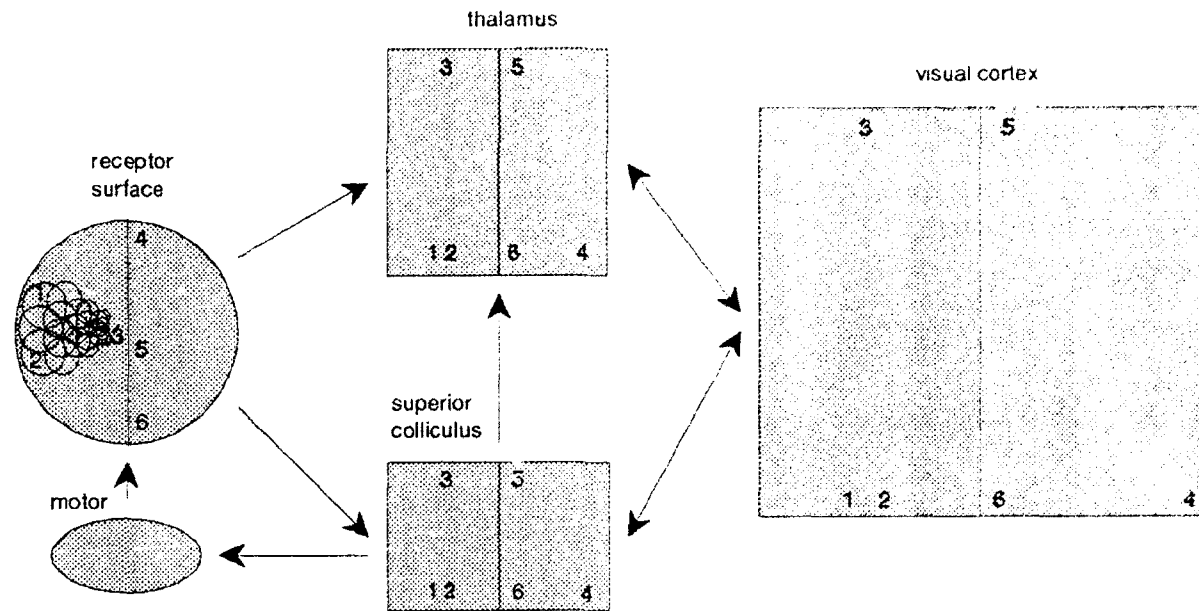
LOG-POLAR MAPPING

The output of the pattern processing subnetwork of the retina is passed to the lateral geniculate nucleus (LGN) of the thalamus. The mapping of the retinal output to the LGN and cortex is not geometrically isomorphic, but can be described by a complex logarithmic transformation (Schwartz, 1980, 1984). First of all, receptors residing on the right half of the retina (figure 1) receive information from the left half of the visual field, and the output ganglion cells project to the right hemisphere, while the right half of the visual field is projected to the left hemisphere. Thus, the visual field, and the retinal surface are essentially split down the vertical midline. Secondly, the large numbers of fibers leaving the foveal region (central area) of the retina are accompanied by a much smaller number of fibers from the periphery that none-the-less represent a much larger region of the visual field. The physical constraints of bundling contributes to the grouping of peripheral fibers on one side of the bundle, because they are displaced by the large number of foveal fibers occupying the other side.

---

[1]Because biological terms are used in this paper to identify different processing structures and layers, some confusion is possible between references to biological data and model constructions. To help prevent this confusion, components of biological structures will be called *cells* and components of artificial structures will be called *elements*.

[2]The appendix provides a list of symbols and abbreviations.

Figure 1. Schematic of the processing stations involved in early vision. Arrows indicate direction of information flow. The numbers represent mappings of retinal locations between the processing stations. The thalamus, superior colliculus, and visual cortex are bilateral structures, divided by the vertical meridian. The small circles on the receptor surface represent receptive fields of individual ganglion output elements showing that the receptive fields increase in size with distance from the center. Similar fields cover the entire surface.

The use of log-polar mapping in the present model is illustrated in figure 1. The computational details of this mapping can be found in Blackburn (1993a). The net result of the mapping places the projections from the central region of the receptor surface at one end of the primary visual cortex, and from the periphery of the receptor surface at the the other end. Foveal magnification on the cortical projection is a consequence of this mapping (Rolls & Cowey, 1970). Only radial lines on the receptor surface will map to vertical lines on the other components, while only concentric arc lines on the receptor surface will map to horizontal lines on the other components. Thus, short vertical or horizontal lines located off the vertical or horizontal meridians of the receptor surface will produce a variety of nonlinear representations on the other maps.

THALAMIC LGN

A comprehensive review of the thalamic lateral geniculate nucleus is available in Jones (1985) from which the following relevant details were extracted. The LGN receives retinotopically the retinal output, and saccade information from the SC. The retinotopic mapping is described by the log-polar transform explained above. Neighborhood relationships are maintained by the mapping. The LGN forwards processed signals to the cortex. The LGN is divided into several laminae, each receiving predominantly a different type of ganglion output cell from the retina. The receptive fields of the LGN principal cells are similar to the receptive fields of the retinal ganglion cells from which they receive input, and the convergence (or divergence) of retinal cells upon thalamic

3

relay cells is quite small.

We have implemented a relatively simple network for LGN processing. The connections for the off-center X pathway are shown in figure 2. The processing in the principal elements (*TX0* and *TX1*) is modulated by inhibitory elements (*IX0* and *IX1*) and suppressed by saccade signals from the SC (Singer, 1977). Principal elements maintain the same receptive fields (0 = off-center, 1 = on-center) as the retinal ganglion elements. If the potential is driven negative, as occurs following a saccade, the output is set to zero.

The potential on the thalamic principal off-center element is given by

$$TX0_{i,j} = max \ \{0, RXO_{i,j} - IX0_{i,j} - S\} \ , \qquad [1]$$

where S is a saccade signal from the superior colliculus and is set to half the maximal output possible from a thalamic principal element when the SC orders a saccade, otherwise = 0.
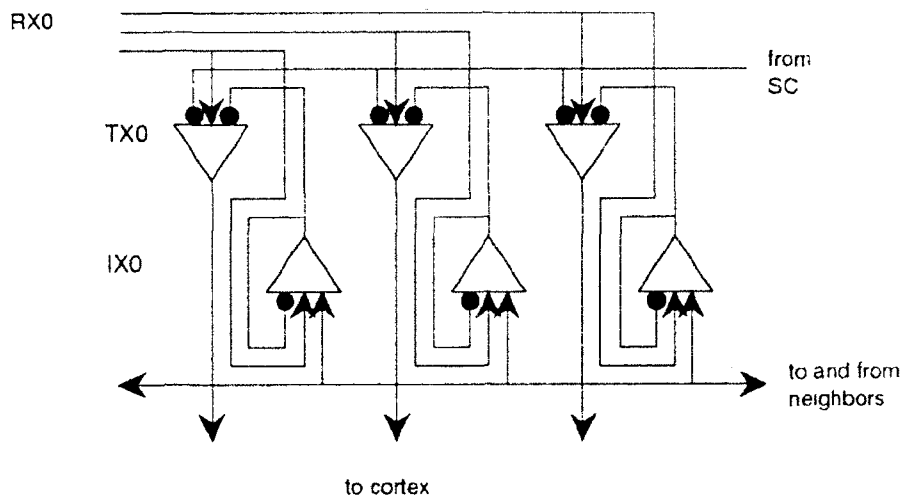
The inhibitory elements attempts to follow the potential arriving from the retina at its principal element. In addition, potentials are distributed laterally within the inhibitory element layer just as they are done in the horizontal layer of the retina (Blackburn, 1993a).

The off-center inhibitory element input buffer (*IX0b*) is given by

$$IX0b_{i,j(t)} = K_1*(RXO_{i,j(t)} - IX0_{i,j(t-1)}) +$$
$$(K_2/6)*\Sigma_{k,l}(IX0_{i+k,j+l(t-1)} - IX0_{i,j(t-1)}) \ , \qquad [2]$$

where RXO is the retinal ganglion element output potential, $IX0$ is the equilibrated inhibitory element potential from the previous time step, k and l define the six nearest neighbors, i and j locate the receptive field center. K_1 and K_2 are constants; in the present examples, $K_1 = 0.5$ and $K_2 = 0.5$.

The sum is normalized by the number of potentially contributing inputs. Normalization is accomplished by using a gain factor for transmission that is inversely proportional to the number of inputs to an element. Biologically, this could reflect the competition for available post synaptic space. The larger the number of inputs from different sources, the less chance that any one source could command a sufficient number of synaptic sites to have a big effect.

4

Figure 2. Model of the on-center pathway in the thalamic LGN. TX0 are principal X type relay elements, IX0 are their inhibitory elements. Connections between lines occur only at "T" junctions in the drawing.

In one time step *(t)*, the potential becomes uniformly distributed in each inhibitory element:
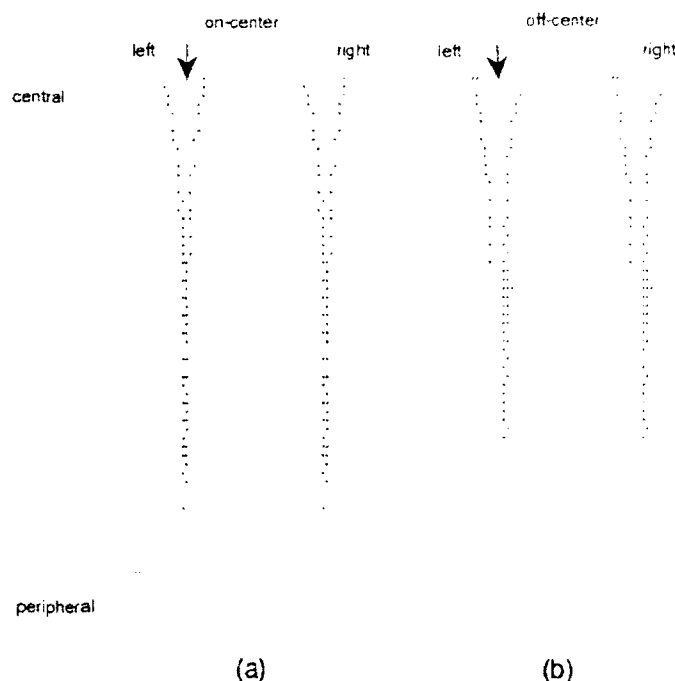
$$IX0_{i,j(t+1)} = IX0b_{i,j(t)} \ . \qquad [3]$$

Similar expressions to [1], [2] and [3] exist for on-center elements *(TX1)*.

The dimensions of the LGN model (and of the cortex model) are essentially set by the retinal projection. In the current implementation, a 64 by 64 array of retinal ganglion elements project to the LGN. Near the center of the retina, there are more ganglion elements than receptors. Thus, a one to many mapping results. A consequence of this overrepresentation of the retinal center is a redundant mass of activity initially in the central region of the projection to the LGN. However, the surround inhibition in the LGN rapidly eliminates this redundant activity, and only the unique edges of activity are passed to the model cortex.

Figure 3 shows the pattern induced in the LGN model by a diagonal line of one pixel in width rotating clockwise 5.6 degrees per program cycle. The division of the pattern near the central projection is due to the redundant representation with surround inhibition that highlights only the contrast borders. The asymmetric off-center activity seen to the right of the on-center activity (centered in column 17 at the arrows) is due to the motion of the pattern to the left on the LGN matrix and the delay in forming and inhibitory surround in the outer plexiform layer of the retina (see Blackburn, 1992a, for a more complete explanation). While the original image width was only one pixel, the broadening of the pattern in the LGN is due to the overlap of receptive fields in the retina.

Figure 3. Activity pattern in the thalamic LGN model induced by a diagonal line of one pixel in width rotating clockwise 5.6 degrees per program cycle on the surface of the retinal model. The activity pattern is moving to the left in the LGN. The left and right bilateral LGN have been placed together. Active on-center elements (with potentials that are greater than zero) are shown in (a) with an asterisk, and off-center elements are shown similarly in (b). Column 17 is indicated in (a) and (b) with arrows.

Processing differs in the LGN model from the retinal model in two important respects: (1) the thalamic inhibitory element is specific to each principal relay element, and (2) there is a one-to-one relationship between input and output elements in the LGN across the projection field. There is, thus, one inhibitory element for an on-center element and a different inhibitory element for an off-center element..

The principal elements of the Y pathway of the LGN (*TY0* and *TY1*) currently accept the output from the motion processing subnetwork of the retina (*RY0* and *RY'*) and are computed in similar fashion to the *TX0* and *TX1* elements.

PRIMARY VISUAL CORTEX

In the mammalian primary visual cortex, feature detectors have been discovered for line orientation, stimulus length and width, movement velocity and direction, spatial frequency, position disparity, and orientation disparity (Orban ,1984).

Feature detectors for lines or edges were divided into simple complex and hypercomplex types by Hubel and Wiesel (1962, 1977). According to their original definitions, simple feature detector cells respond maximally to a line with a specific orientation and location in a very circumscribed region of the visual field. The complex feature detector cells respond to a line of specific orientation but allow a larger region of the visual space. The complex cells do not permit the width of the stimulus line to increase proportionately. Hypercomplex cells are similar to complex cells except that they can be inhibited by extending the length of the line beyond the excitatory receptive field.

Definitions of simple and complex cells were modified when more complex stimuli were used to test the cells receptive field characteristics. Drifting sinusoidal or square wave gratings were used to stimulate the entire receptive field of a cortical cell at one time. Simple feature detectors were then defined as cells that responded in a linear fashion (defined by the temporal modulation of the cell's firing rate by the rate of drift of the grating interacting with its spatial frequency) to a sinusoidal grating with a spatial periodicity that matched the periodicity of the center and the surround of the cell's receptive field. Complex feature detectors were cells that responded in a nonlinear fashion (Maffei and Fiorentini, 1973). The stimulus conditions associated with the maximal firing rate of the cell defined the cell's preferred orientation and spatial frequency.

Within a few degrees of the visual field center, the receptive fields of cortical cells approximate the receptive fields of single retinal ganglion cells except that they are more sharply tuned for spatial frequency (Maffei and Fiorentini, 1973). When a stimulus grating pattern is in phase with retinal ganglion cell's receptive field, the cell will respond maximally, but when the pattern is 180 degrees out of phase, the retinal ganglion cell's activity will be suppressed. Because the sinusoidal grating covered the entire receptive field of a cell, a linear response indicated an integration of a single pair (or several pairs of nonoverlapping) inhibitory and excitatory zones, while a nonlinear response indicated an integration of multiple overlapping inhibitory and excitatory zones (a laterally drifting grating across the receptive field of a complex element could generate a continuous response). The size of the inhibitory and excitatory zones contributed to the preferred spatial frequency of the simple or complex cell. If the grating frequency was greater than the width of the zones, then both zones would be coactivated by a stationary or drifting grating regardless of phase and the cells response would be reduced. Low spatial frequencies of square wave gratings could also coactivate both zones, but would appear to the cell as an edge, or border between bright and dark regions, moving across the field. The distinct classification of hypercomplex cells has lost favor due to the observation that both simple and complex cells can show end-stop inhibition (Dreher, 1972; Orban, 1984).

Motion Selectivity

Many simple and complex cells showed a preference for the direction of drift of the oriented line or grating. The preferred direction was always orthogonal to the preferred line orientation (Schiller et al., 1976). In addition, some complex cells respond preferentially to stimulus velocity, that is - a stimulus moving in a particular direction with a particular speed (Sekuler et al., 1990). Simple cells do not generally have velocity sensitivity, except that they respond best to slowly moving stimuli (Maffei and Fiorentini, 1973). The rather complete dependence of simple cells on input from the X pathway, which is a low pass system, may account for their inability to respond to high stimulus velocities. The complex cells, on the other hand, receive input also from the Y pathway, which is a high pass system and may give the complex cells the capability to respond to faster moving events.

Direction of Contrast Independence

At some point in image processing, a line must have an effect that is independent of whether it is a bright line on a dark background, or a dark line on a bright background. At that point, the "lineness" of the feature is direction-of-contrast independent. While a common house fly prefers a dark target on a bright background (Wehner, 1981), humans have little difficulty reading bright text on a dark background, nor the reverse, thus demonstrating direction-of-contrast independence.

There is evidence for the early confluence of on-center and off-center activity in the primary visual cortex. Bishop et al. (1971) reported that a majority of line oriented simple cells responded to both edges of a moving narrow bar of light on a dark background. Schiller (1982) found that chemical blocking of the retinal on-center mechanisms eliminated the light-edge response of cortical cells but "had no discernible effect on orientation and direction specificities". Their data indicated that both on-center and off-center activity converge on the same line orientation detectors. A similar finding was reported by Sherk and Horton (1984).

The early confluence of on-center and off-center activity would save duplicate processing at later stages and allow for the general definitions of features independent of its contrast relationship to the background. In the current implementation (figures 4 to 6), a direction-of-contrast independence of the oriented line detectors is achieved by the projection of both on and off-center thalamic elements with the same receptive fields to the same cortical elements .

Line Orientation Selectivity

The process actually starts in the interaction of receptors, horizontals, and bipolars in the retina. Only one of a complement pair of on-center and off-center bipolars that have the same receptive field center can be active at any moment, yet the surrounding neighborhood of an active bipolar of one type is usually composed of active bipolars of the other type. Therefore, when the cortical input layer elements receive a fuzzy (slightly

8

defocussed) distribution of on-center and off-center activity from the retina, the neighboring projections of contrasting types can reinforce the pattern center.

Thalamic input to the cortex is distributed primarily to stellate cells in layers 4A and 4C. The slower conducting, sustained X-type activity is projected to cortical layers 4A and 4Cβ, while the faster conducting, transient Y-type activity is projected to cortical layer 4Cα. Layer 4C cells then project to other cortical layers. There is ample evidence that the other cortical layers receive direct input from the LGN as well. In the present paper, we have identified processing elements by using labels taken from the cortical layers in which we speculate that the modeled biological activity takes place.

In order to be consistent with the finding of Creutzfeldt and Ito (1968) that the excitatory receptive fields of cortical elements are circular, we distributed the thalamic input locally in a small field about each target cortical element. In this way, each cortical element receives excitatory input from a similar number of on-center and off-center thalamic elements covering the same receptive fields (figures 4 and 5). Neighboring cortical elements receive similar thalamic input offset by one thalamic receptive field from the next. This divergence and overlap in the projection defocusses or fuzzifies the image, but allows for the definition of oriented line segments of length equal to the degree of fuzziness. A longer radius of integration would increase fuzziness of the projection and increase the potential line orientation resolution.

Because intracortical inhibition has been shown to participate in the definition of orientation sensitivity (Sillito et al., 1980), we used inhibition to shape the directional sensitivities of the cortical elements similar to the proposal of Heggelund (1981). Additionally, Matsubara et al. (1985) found that simple cells formed inhibitory connections preferentially between detectors with orthogonal orientations.

The 4Cβ elements and the 4B line orientation detector elements receive, with a small convergence equal to the divergence of the thalamic input, excitatory direct input from thalamic elements (figure 4). The 4B elements then receive indirect thalamic inhibition via the 4Cβ elements from a specific region of their surround. The extent of the inhibition in one direction is set at twice the radius of the integrated input from the thalamus. Each 4B element receives inhibition from 4Cβ elements only along one diameter (figure 5). The directionally oriented inhibition could be accomplished in natural systems by a linear arborization of the 4B cell's basal dendrites (Tieman and Hirsch, 1982), or by horizontally aligned stellate interneurons (Braak, 1980). The result of the linear accumulation of inhibition from the 4Cβ elements is an orientation preference in the 4B element that is normal to its inhibitory input. In addition, the 4B oriented feature detector receives inhibitory input from another detector that has the same receptive field but has an orientation preference normal to its own orientation preference (figure 6). The output of the 4B feature detector is half-wave rectified, passing only positive activity.
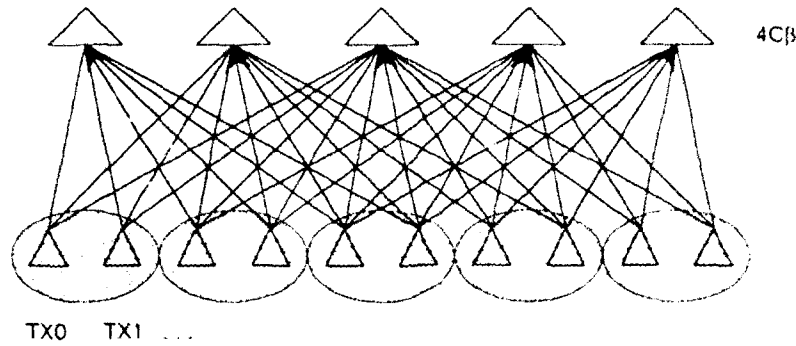
9

Figure 4. Confluence of pairs of off-center and on-center thalamic input and their distribution to layer *4Cb* elements. The distribution pattern occurs over two dimensions, though only one dimension is shown in the figure. A similar distribution of thalamic X activity is also made to layer *4B* elements.

In the present model, it is thus the direction of the lateral spread of inhibition that defines the preferred orientation tuning of the cortical element, and not any specificity in the thalamic projection as was suggested by Hubel and Wiesel (1972). Lateral inhibition that defines the line orientation also removes the fuzziness due to the defocussed thalamic projection.

The input to the layer $4C\beta$ element is given by

$$4C\beta_{i,j} = \Sigma_{k,l} \; \lambda * (TX0_{i+k,j+l} + TX1_{i+k,j+l}). \tag{4}$$

where k and l index the convergence of thalamic input to the layer $4C\beta$ element and $\lambda = (k*k + l*l)^{-1/2}$. Currently, the degree of convergence (and divergence) of thalamic input to layer $4C\beta$ is 9 to 1. That is, a layer $4C\beta$ element collects input from thalamic projections that map within a distance of +/- 1 element from its center. This degree of convergence could be increased in a system with higher resolution (a larger number of receptive fields with relatively smaller receptive field sizes).
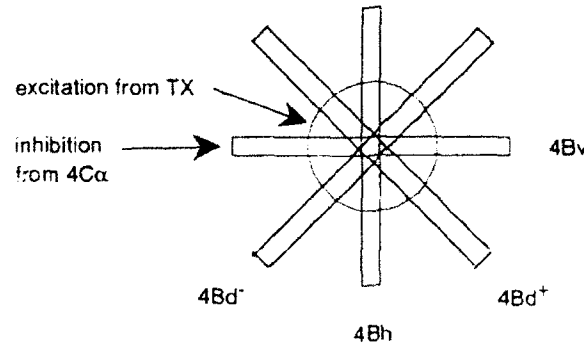

Figure 5. Orientation of inhibitory input to layer *4B* simple feature detectors (shown in two dimensions). The detectors are labeled by their orientation sensitivities rather than the orientation of the inhibition that they receive. The four orientation detectors shown have receptive field centers essentially in the same location on the receptor surface.

The potential on the oriented detectors in layer 4B is set as

$$4Bv_{i,j(t)} = max \{0, \Sigma_{k,l} \lambda^*(TX0_{i+k,j+l} + TX1_{i+k,j+l})(t) - (\gamma^*\Sigma_m 4C\beta_{i,j+m(t-1)} + 4Bh_{i,j(t-1)})\}$$

$$4Bh_{i,j(t)} = max \{0, \Sigma_{k,l} \lambda^*(TX0_{i+k,j+l} + TX1_{i+k,j+l})(t) - (\gamma^*\Sigma_m 4C\beta_{i+m,j(t-1)} + 4Bv_{i,j(t-1)})\}$$

$$4Bd^-_{i,j(t)} = max \{0, \Sigma_{k,l} \lambda^*(TX0_{i+k,j+l} + TX1_{i+k,j+l})(t) - (\gamma^*\Sigma_m 4C\beta_{i-k,j+k(t-1)} + 4Bd^+_{i,j(t-1)})\}$$

$$4Bd^+_{i,j(t)} = max \{0, \Sigma_{k,l} \lambda^*(TX0_{i+k,j+l} + TX1_{i+k,j+l})(t) - (\gamma^*\Sigma_m 4C\beta_{i+m,j+m(t-1)} + 4Bd^-_{i,j(t-1)})\}, \quad [5]$$

where $\gamma = k_{max}^{-1}$; λ, k and l are defined as in [4]; $m = 2^*\lambda$ is the distance over which inhibition is integrated; and h identifies the horizontal orientation preference, v the vertical orientation preference, and d^+ and d^- the diagonal orientation preferences relative to the surface of the cortical map.
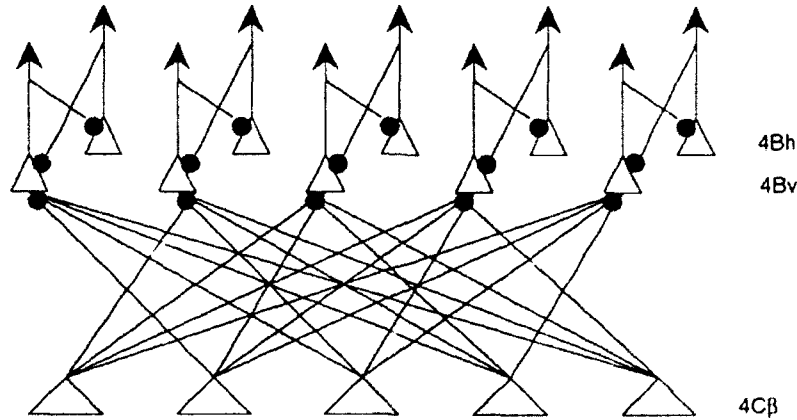

Figure 6. One-dimensional horizontal distribution of inhibition from 4Cb elements to 4Bv elements that contribute to their sensitivity to vertical lines. Feedback inhibition from horizontal line detectors (4Bh) that have received inhibition only along the vertical axis completes the process. When a horizontal line is present in the input, both 4Bv and 4Bh detectors receive the same amount of excitation from the thalamus, but the greater inhibition from the series of 4Cb elements along the projection of the horizontal line stimulus onto the 4Bv elements allow the 4Bh detectors to overpower the complementary 4Bv detectors and represent the stimulus feature.
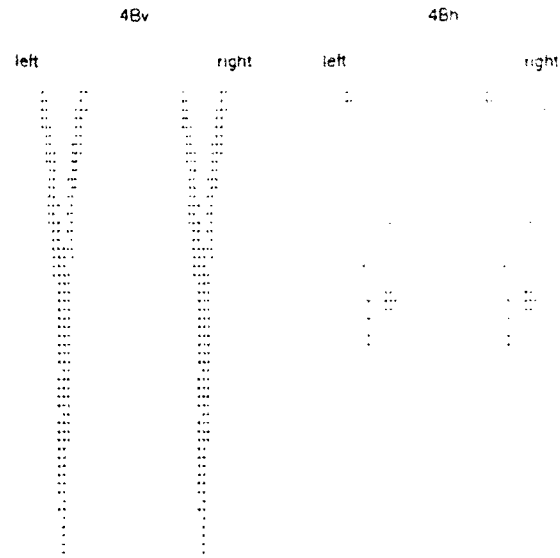
The activity patterns produced in the layer *4B* line orientation detectors by the pattern of activity shown for the LGN in figure 3 are given in figures 7 a-d. Since the LGN is forwarding essentially a pair of vertical lines to the cortex, the *4Bv* detectors in the majority represent the pattern. In the central projection, however, the line branches in a "Y" formation, thus evoking activity in the diagonal line detectors. Only sporadic activity in the horizontal detectors is seen.

Sectorization of the Cortex

The model cortical surface is organized into overlapping sectors (figure 8) that roughly model the concept of the hypercolumn. In the present version, each sector is composed of 8*8 individual receptive fields. The sectors overlap their neighbors by 25%. Motion is analyzed independently in each of the overlapping sectors. The direction of motion on the horizontal or vertical surface of the cortex is determined. The motion information is used to determine activations of direction selective complex elements.
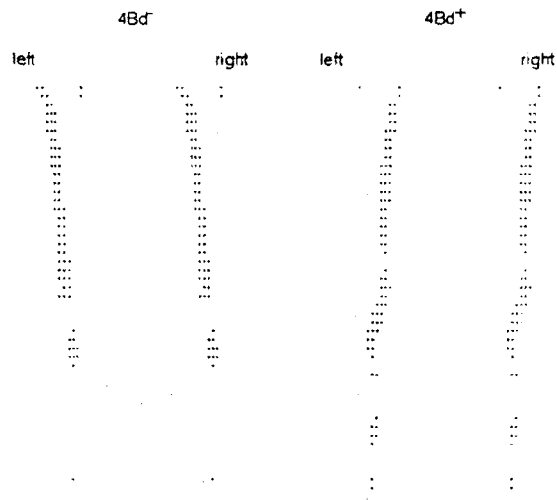
Motion Analysis In The Cortex

The method of motion analysis that we use is that of feed-forward lateral inhibition (Blackburn et al., 1987). In Blackburn et al. (1987), motion was analyzed in the inner plexiform layer of the retina model. In the present implementation, the essentials of the earlier motion analysis are maintained but the site of the processing has been moved centrally to layers *4Cα* and *5A* of the cortex. The input from the Y pathway of the LGN is distributed to four sheets of elements in layer *4Cα* in such a way that the activity is slightly offset laterally from a direct topographic mapping of the input. The *4Cα* elements then inhibit the underlying pyramidal elements in layer *5A*. For a receptive field of arbitrary size, we have found it practical to assign only four directionally sensitive pyramidal elements in each sector, each one sensitive to movement in one of the four cardinal directions relative to the cortical surface. Each of the four sheets of *4Cα* elements receives input from a different cardinal direction. Movement of a spot of light on the receptor surface will produce before it, on one or another *4Cα* sheet, a moving spot of inhibition that opposes conduction of the *4Cα* potential to the directionally sensitive pyramidal element in layer *5A*. This defines the null response direction of movement in the receptive field for that pyramidal element. Diagonal motion on the cortical surface is measured by the joint activity of orthogonal pairs of the four cardinal direction motion detectors.

12

Figure 7. Responses of simple line orientation detectors to input from the LGN pattern represented in figure 3.
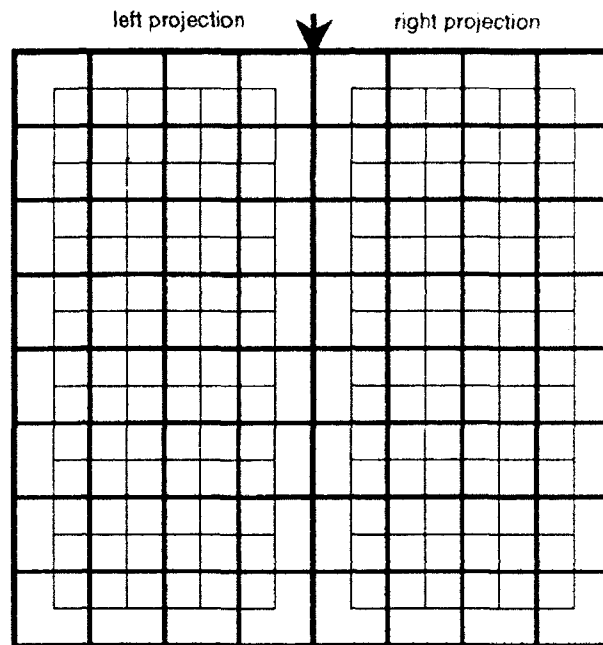
left projection   right projection

Figure 8. Sectorization of the cortical plane. The plane is first divided down the vertical midline (arrow). The left half of the cortical plane receives projections from the left half of the receptor surface (via the thalamic LGN), and the right cortical plane receives projections from the right half of the receptor surface. There are 53 sectors on each side, each the size of the medium line density squares, and each overlapping each of its neighbors by 25% (the size of the smallest squares).

Complex Elements

Complex elements are subject to line orientation, spatial period, and direction of motion. We insure that motion information is available locally from the LGN Y pathway, analyzed between layers $4C\alpha$ and $5A$. The diameters of the receptive fields of the Y retinal ganglion cells are approximately 3 times larger than the diameters of X retinal ganglion cells at all eccentricities (Perry, Oehler, and Cowey, 1984). While the receptive field sizes of our model Y and X retinal elements are the same, we have, for convenience, integrated Y element output in 8 by 8 element sectors that approximate the larger receptive fields of Y ganglion cells. Motion from the retinal Y elements and spatial periodicity of X element activity is computed over the size of the integrated Y element sectors. The range and resolution of the spatial periods that can be discriminated within a sector are limited by the size of the receptive fields and their number within the sector.

14

In the present implementation of complex elements, the input from the simple detectors of layer *4B*, with a short delay, determines the sensitivity to line orientation and spatial period. Direction specificity of the *3B* element is due to a gating input from *5A*. If motion in a particular direction is indicated, and if there is activity from a feature detector located in the direction of the origin of the motion, then the activation (if any) of a like type feature detector located at a given distance (the spatial period sensitivity) in the direction of motion is passed to the complex element. The circuit diagram for the input to a complex element with sensitivity to spatial periods of two receptive fields, with a horizontal orientation, and with a movement preference to the left, is given in figure 9.
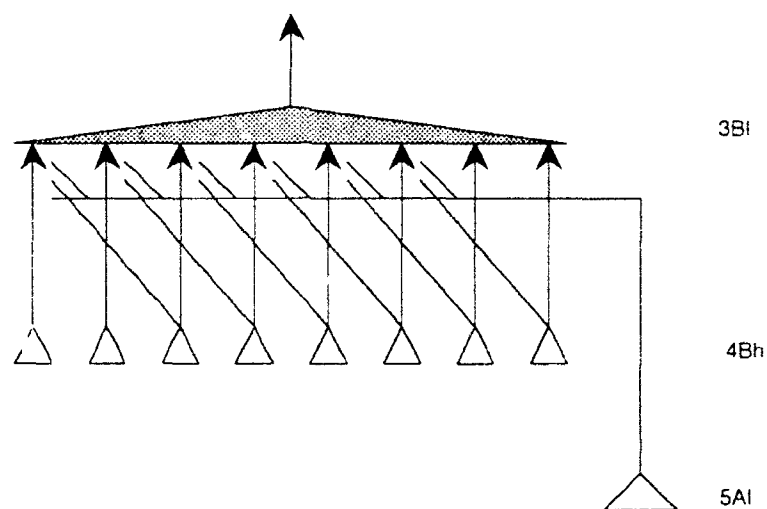

Figure 9. The layer *3B* complex element is selective to orientation by virtue of its input from an orientation selective element in layer *4B*, to direction of motion due to the gating action of a directionally selective element in layer *5A*, and to spatial period due to the feed forward lateral facilitation of layer *4B* elements with a given delay (distance) before termination upon the *3B* element. Increased periods are achieved by increased delays.

Complex elements are designed to respond preferentially to directed motion orthogonal to their preferred line orientation. Since we have simple elements that respond preferentially to four different orientations (two diagonals, the vertical and horizontal), we can construct complex elements with eight different combinations of direction and orientation preference. The size of the sectors determines the degree of location uncertainty of events within a sector. The spatial frequencies available for analysis are determined by the lateral offset of the *4B* contributions. We have limited these to origins within neighboring sectors.

Within each sector, the orientation and direction of motion selectivity are accomplished by combining in a logical 'and' element, activities from oriented simple line detectors and motion detectors with direction preferences orthogonal to the line orientation. When the 'and' element becomes active, activity from a line detector with the same orientation and with a specific spatial offset from the former line detector in the direction of motion is passed to the complex element. All such pairs with the same spatial period selectivity

15

are summed within a sector. When the dimensions of the sector's receptive field is 8*8, the complex elements are selective for spatial periods or offsets of 2,3,4 and 5 simple detector elements. The potential on a complex element is given by

$$3Br_{k(t)} = \Sigma_p \ (4Bv_{i,j(t)} \ / \ 5Ar_{(t)} \ \& \ 4Bv_{i,j-k(t-1)})$$

$$3Bl_{k(t)} = \Sigma_p \ (4Bv_{i,j(t)} \ / \ 5Al_{(t)} \ \& \ 4Bv_{i,j+k(t-1)})$$

$$3Bd_{k(t)} = \Sigma_p \ (4Bh_{i,j(t)} \ / \ 5Ad_{(t)} \ \& \ 4Bh_{i-k,j(t-1)})$$

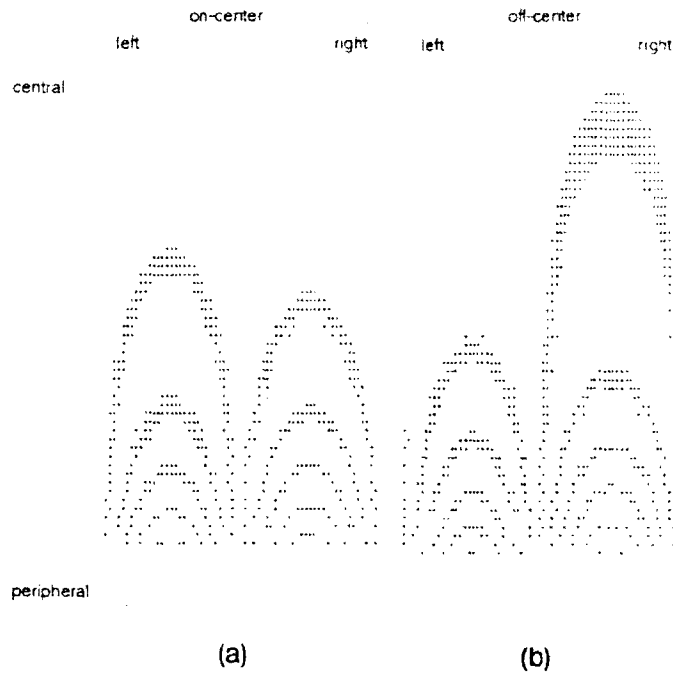$$3Bu_{k(t)} = \Sigma_p (4Bh_{i,j(t)} \ / \ 5Au_{(t)} \ \& \ 4Bh_{i+k,j(t-1)}), \qquad [6]$$

where k indexes the offset that defines the spatial period; p represents the number of pairs of oriented detectors k receptive fields apart within the sector; r is a right motion detector. l is a left motion detector, d is a down motion detector, u is an up motion detector; v is a vertical orientation detector, h is an horizontal motion detector.

Diagonal complex elements ($3Bru_k$, $3Bld_k$, $3Brd_k$, $3Blu_k$) are computed similarly except that the direction of motion is determined by two $5A$ elements, and the search vector is determined by adjustments to both i and j indices.

$$3Bru_{k(t)} =$$
$$\Sigma_p \ (4Bd_{i,j(t)} \ / \ 5Ar_{(t)} \ \& \ 5Au_{(t)} \ \& \ 4Bd_{i+k,j+k(t-1)}).$$
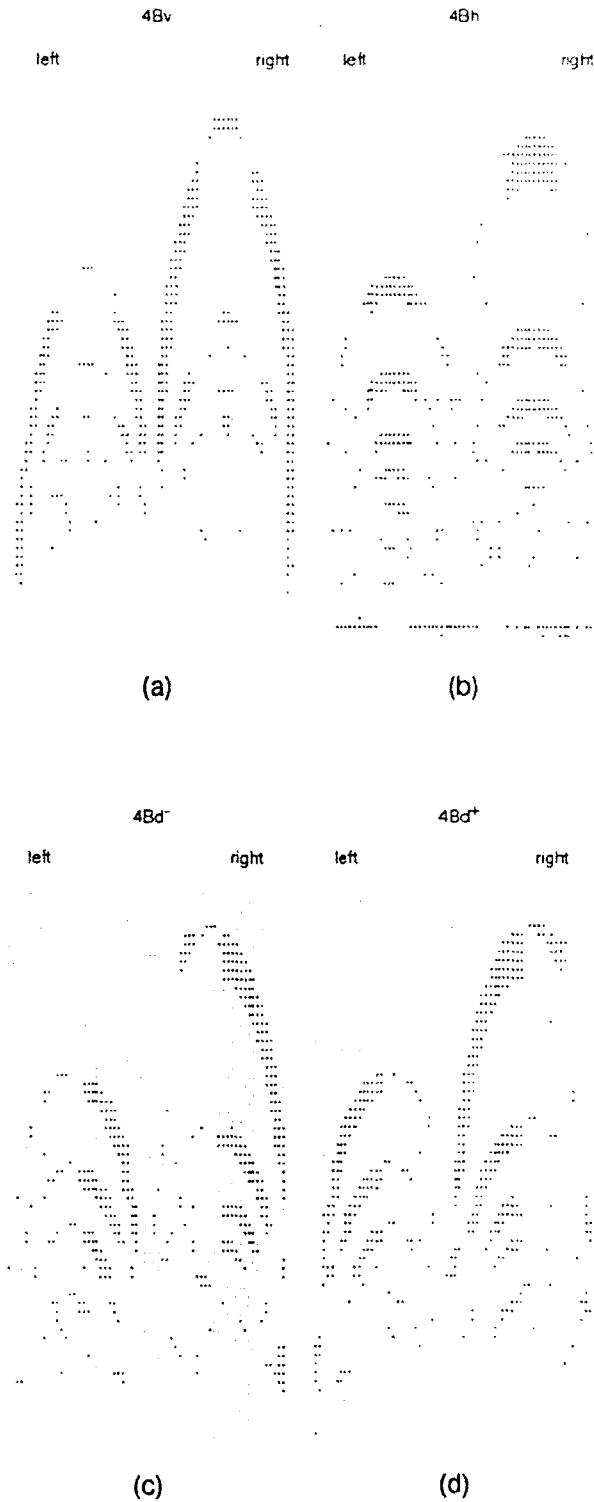$$[7]$$

BEHAVIOR

The responses of elements within the system were tested by using stimulus grating patterns of different spatial frequency that were drifted across the entire visual field. Figures 10 (a) and (b) show the activity patterns in the on-center and off-center principal elements of the model LGN to a square wave grating pattern with a period of 15 pixels. The rate of drift was one pixel per program cycle. Nonzero activity is indicated in the figures by an asterisk. The effects of the log-polar mapping are evident in the nonlinear projection of the regular grid pattern to the thalamic matrices. Differences between the patterns of activity that developed in the LGN of figure 10 are due to the phase difference of the grid lying on the two hemiretinae. The absence of a pattern in the far peripheral projection is due to the large receptive fields relative to the size of the grid and the effects of surround inhibition. The central projection also fails to respond to the grid, but because the grid lines are too thick to generate any contrast between the center and the surrounding receptive fields. However, the off-center elements do respond to the advance of the dark edge.

16

Figure 10. Projection of a grating pattern upon the LGN model.

Figures 11 (a-d) show the activity patterns of the four different types of simple feature detectors of the model cortical layer *4B* in response to the patterns of activity in the LGN of figure 10. It can be seen that while the detectors respond locally to the orientation of their input, they also respond as a group of like type to represent the global features of the thalamic pattern.

The activity profiles of representative layer *3B* elements taken from different sectors during tests with different grid spatial frequencies are given in figure 12. The sampled elements were always the detectors for the lowest spatial frequency in each sector (5 overlapping receptive fields). These profiles are similar to those recorded from biological complex cells in that activity was decreased by greater than 50% for each octave change in spatial frequency (Maffei and Fiorentini, 1973). Elements located in sectors that received the peripheral projection were more responsive to grating patterns with lower spatial frequencies. High spatial frequency gratings did not activate the complex elements in the peripheral projection due to their absence in the thalamic projection (figure 10). The opposite was true for elements receiving the central projection. This sensitivity is directly a function of the receptive field sizes of the X retinal ganglion elements because all sectors of the cortex model were processed identically.

17

Figure 11. Activity patterns in layers of simple feature detectors responding to input from the pattern present in the LGN of figure 10: vertical line detectors (a), horizontal line detectors (b), diagonal line detectors for -45 degrees (c), and for +45 degrees (d).
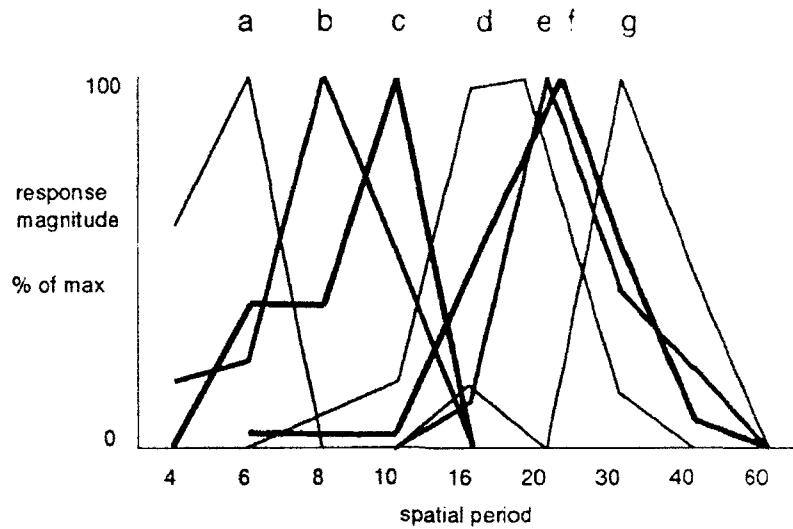
18

Figure 12. Activity profiles of complex elements taken from locations that receive projections with increasing eccentricity (a-g). Note that the spatial period scale is nonlinear.

The activity profiles of figure 13 show that sampling within a sector while changing the spatial frequency of the grating pattern also resulted in the discrimination of the pattern by differently tuned detectors, though sensitivities varied over a much smaller range than was the case between sectors. The differences in tuning of the spatial frequency detectors within sectors was due to the connectivity patterns that passed lateral facilitation over different distances (figure 9).
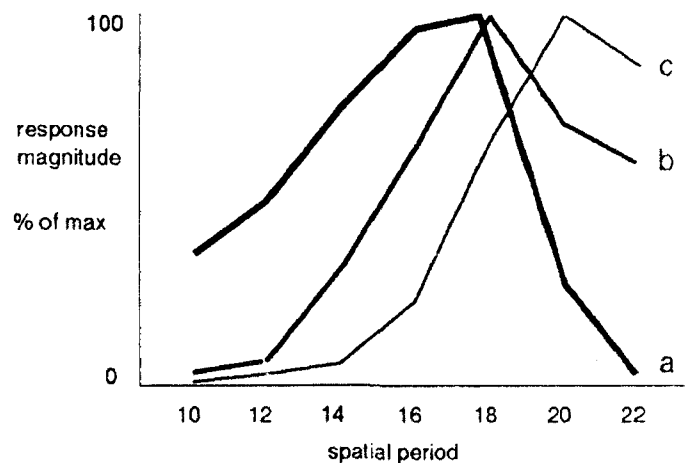

Figure 13. Activity profiles of three complex elements located in the same paracentral sector and representing the same orientation and direction selectivities but with increasing spatial period sensitivities from a to c.

19

DISCUSSION

The dependence of orientation preference in simple line detector elements on a dendritic field bias has received checkered confirmation in the literature. Tieman and Hirsch (1982) found that kittens deprived of lines of particular orientations developed abnormal distributions of dendritic field orientations. Interestingly, the abnormalities were in orientations of fields that were orthogonal to the expected projections of the missing stimulus orientations - a finding compatible with the mechanisms employed in the present application. However, Martin and Whitteridge (1984), while able to measure a dendritic bias in all cells studied, found no consistent relationship between the dendritic field bias and line orientation preference. The negative results of Martin and Whitteridge (1984) could have been due to their sampling from several types of neurons from cortical layers II through VI, whereas the positive results of Tieman and Hirsch (1982) could have been related to their limited sampling of pyramidal cells in layer III. No such distribution abnormalities were seen in layer IV stellates taken from the same preparations. Both studies faced the same problem of estimating the projection pattern of the thalamic input, which due to the log-polar mapping, is not straight forward and could complicate efforts to find correlations with dendritic field orientations.

As an alternative to the use of oriented dendritic fields for the organization of orientation sensitivity, Vidyasagar (1987) argued that slight asymmetries in the LGN responses to oriented stimulus grids of different spatial frequency could contribute to the sharp orientation tuning of cortical simple elements if LGN elements with orthogonally asymmetric responses would converge in parallel to a pair of cortical elements, one of which was inhibitory. In order to implement a Vidyasagar model, we would have to account for the response asymmetries in the LGN or retina, and provide a way to match orthogonal asymmetries in the input to cortical elements and their associated inhibitory interneurons. There is also a question of whether the asymmetries in receptive fields from the thalamus could provide a complete and systematic representation of orientations within each region of the visual field.

The early convergence of on-center and off-center activity upon the simple feature detectors would appear to be at odds with the observation that some simple cells have distinct center-surround receptive fields (Movshon et al, 1978), but other papers indicate that many simple cells receive both on-center and off-center thalamic input (Bullier et al., 1982; Orban, 1984). The present model will produce the inhibitory effects of the surround by lateral inhibition at three levels. First, a broadening of illumination (or of darkness) will reduce the output of the retinal ganglion elements. Second, the surround inhibition of the thalamic LGN attempts to further sharpen the contrast region. Third, the mutual inhibition between orthogonal line detectors in the cortex will reduce the response of a line detector if activity is increased on its sides. Simple feature detectors of the present model, however, do not respond differently to a bright line that is later replaced exactly by a dark line.

In the monkey, radial and tangential orientations (relative to the foveal center) predominate (Bauer and Dow, 1989). Because of the log-polar mapping of retinal output

20

to the cortex, longitudinally oriented activity on the cortex translates to radially oriented activity on the retina, while latitudinally oriented activity on the cortex translates to tangentially or concentrically oriented retinal activity. This greatly simplifies the local computations required in the cortex model to produce these radial and concentric orientations relative to the visual field, for the integrations can be performed either in parallel or in orthogonal directions across the cortical surface.

A non-parallel distribution of radial dendritic fields (relative to the cell body) would provide for a variety of orientation sensitivities that have been described by Hubel and Wiesel (1972) and others. The systematic shift in orientation sensitivity found with an electrode probing tangentially to the cortical surface could be a function of the realities of packing densities. The dendritic processes of several cells cannot occupy exactly the same space in the plexus, yet the proximity of their cell bodies allows them to share similar receptive fields and to send their dendrites into neighboring space, offset by the few degrees width of the radial dendritic field.

The meaning of activity in a complex element can be deduced from the several factors that contribute to that activity. These factors are directional movement and spatially distributed line features oriented orthogonal to the direction of movement. Thus, an active complex element indicates that some line segments with a particular spatial separation are moving together in a direction orthogonal to their orientation, that is, complex elements indicate the motion of a bar of some width. The temporal delay in transmission from the facilitating *4B* element to the *3B* element is currently independent of distance or spatial period. This can yield a type of velocity detector for a single oriented line moving with sufficient speed to activate both *4B* elements of a pair within the sort period of the delay. Complex elements that are sensitive to greater spatial periods would also be sensitive to greater stimulus velocities. If biological complex cells processed information similar to the algorithm offered herein, then a correlation would be expected between velocity sensitivity to a single moving edge or line and spatial period sensitivity. Cells that responded better to larger spatial periods should also respond better to faster moving lines.

The dependence of the complex elements on the activity of simple elements may be a requirement that is avoided by biological systems. Evidence, summarized by Wilson et al. (1990), suggests that the complex cells are organized in parallel with the simple cells. Yet other evidence that the component contributing receptive fields of a complex cell behave like those of simple cells (Movshon et al., 1978; Heggelund, 1985) is a compelling argument for the conservative use of the simple element's output to provide orientation selectivity for the complex element.

The complex elements provide information that may be predictive. Because both spatial frequency and direction participate in the activation of a complex element, the complex element can predict or anticipate the next location (or appearance) of a moving feature. This capability could be useful in solving problems in perceptual invariance under dynamic conditions (which are by far the most commonly met with in the real world of a behaving system).

REFERENCES

Bauer, R., and B.M. Dow. (1989). "Complementary global maps for orientation coding in upper and lower layers of the monkey's foveal striate cortex." *Experimental Brain Research,* 76, pp. 503-509.

Bishop, P.O., J.S. Coombs, and G.A. Henry. (1971). "Responses to visual contours: Spatio-temporal aspects of excitation in the receptive fields of simple striate neurones." *Journal of Physiology,* 219, pp. 625-657.

Blackburn, M.R. (1990). "A general architecture for classical conditioning of perceptual-motor sequences." *IJCNN 90,* II, pp. 293-298.

Blackburn, M.R. (1993a). "A simple computational model of center-surround receptive fields in the retina." TD 2454, NCCOSC RDT&E Division, San Diego, California.

Blackburn, M.R. (1993b). "Machine visual targeting modeled on biological reflexes." TD 2455, NCCOSC RDT&E Division, San Diego, California.

Blackburn, M.R., and H.G. Nguyen. (1989). "Biological model of vision for an artificial system that learns to perceive its environment." *IJCNN Proceedings,* vol. II, pp. 219-226.

Blackburn, M.E., and H.G. Nguyen.(1990). "Modeling the biological mechanisms of vision: Scan paths." In X. Avula (ed.) *Mathematical and Computer Modeling in Science and Technology.* Oxford: Pergamon Press, pp. 311-316.

Blackburn, M.R., H.G. Nguyen, and P.K. Kaomea. (1987). "Machine visual motion detection modeled on vertebrate retina." *SPIE Proceedings,* vol. 980, pp. 90-98.

Braak, H. (1980) "Architectonics of the human telecephalic cortex." *Studies in Brain Function,* vol. 4, Berlin: Springer-Verlag.

Bullier, J., M.J. Mustari, and G.H.Henry. (1982). "Transformation of receptive field properties of LGn neurons and first order S cells in cat striate cortex." *Journal of Physiology,* vol. 47, pp. 417-438.

Creutzfeldt, O.D., and M. Ito. (1968). "Functional synaptic organization of primary visual cortex neurons in the cat." *Experimental Brain Research,* 6, pp. 324-352.

Dowling, J.E. (1987). *The Retina: An Approachable Part of the Brain.* Cambridge, MA: Belknap Press.

Dreher, B. (1972). "Hypercomplex cells in the cat's striate cortex." *Investigative Ophthalmology,* vol. 11, pp. 355-356.

Ewert, J.-P. (1987). "Neuroethology of releasing mechanisms: prey-catching in toads." *Behavioral and Brain Sciences,* vol. 10, pp. 337-405.

Heggelund, P. (1981). "Receptive field organization of simple cells in cat striate cortex." *Experimental Brain Research*, 42, pp. 89-98.

Heggelund, P. (1985). "Receptive field organization of simple and complex cells." In Rose, D. and Dobson, V.G. (eds.) *Models of the Visual Cortex*. New York: Wiley, pp. 358-365.

Hubel, D.H., and T.N. Wiesel. (1962). "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex." *Journal of Physiology*, 160, pp. 106-154.

Hubel, D.H., and T.N. Wiesel. (1977). "Functional architecture of macaque monkey visual cortex." *Proceedings of the Royal Society of London, B.*, 198, pp. 1-59.

Jones, E.G. (1985). *The Thalamus*. New York: Plenum Press.

Maffei, L., and A. Fiorentini. (1973)."The visual cortex as a spatial frequency analyser." *Vision Research*, 13, pp. 1255-1268.

Martin, K.A.C., and D. Whitteridge. (1984). "The relationship of receptive field properties to the dendritic shape of neurones in the cat striate cortex." *Journal of Physiology*, 356, pp. 291-302.

Matsubara, J., M. Cynader, N.V. Swindale, and M.P. Stryker. (1985)."Intrinsic projections within visual cortex: Evidence for orientation-specific local connections." *Proceedings of the National Academy of Sciences (USA)*, 82, pp. 935-939.

Movshon, J.A, I.D. Thompson, and D.J. Tolhurst. (1978). "Spatial summation in the receptive fields of simple cells in the cat's striate cortex." *Journal of Physiology*, 283, pp. 53-77.

Orban, G.A. (1984). *Neuronal Operations in the Visual Cortex. Studies in Brain Function*, vol. 11. Berlin: Springer-Verlag.

Rodieck, R.W. (1965). "Quantitative analysis of cat retinal ganglion cell response to visual stimuli." *Vision Research*, 5, pp. 583-601.

Rolls, E.T., and A. Cowey. (1970). "Topography of the retina and striate cortex and its relationship to visual acuity in rhesus monkey." *Experimental Brain Research*, 10, pp. 298-310.

Ruff, P.I., J.P. Rauschecker, and G. Palm. (1987). "A model of direction-selective "simple" cells in the visual cortex based on inhibition asymmetry." *Biological Cybernetics*, 57, pp.147-157.

Schiller, P.H. (1982). "Central connections of the retinal On and off pathways." *Nature*, 297, pp. 580-583.

Schiller, P.H., B.L. Finlay, and S.F.Volman. (1976). "Quantitative studies of single cell properties in monkey striate cortex. I-V." *Journal of Neurophysiology*, 39, pp. 1288-1374.

Schwartz, E.L. (1980). "A quantitative model of the functional architecture of human striate cortex with application to visual illusion and cortical texture analysis. *Biological Cybernetics*, 37, pp. 63-76

Schwartz, E.L. (1984). Anatomical and physiological correlates of visual computation from striate to inferotemporal cortex." *IEEE Trans Systems, Man and Cybernetics*. SMC-14, pp. 257-271.

Sekuler, R., S. Anstis, O.J. Braddick, T. Brandt, J.A. Movshon, and G. Orban. (1990). The perception of motion. In L. Spillmann and J.S. Werner (eds.). *Visual Perception: The Neurophysiological Foundations*. San Diego: Academic Press, pp. 205-230.

Sherk, H., and J.C.Horton. (1984). "Receptive field properties in the cat's area 17 in the advance of on-center geniculate input." *Journal of Neuroscience*, 4, pp. 381-393.

Sillito, A.M., J.A. Kemp, J.A. Milson, and N.Berardi. (1980). "A reevaluation of the mechanisms underlying simple cell orientation selectivity." *Brain Research*, 194, pp. 517-520.

Singer, W. (1977). "Control of thalamic transmission by corticofugal and ascending reticular pathways in the visual system." *Physiological Reviews*, vol. 57, pp. 386-420.

Stone, J., B. Dreher, and A. Leventhal. (1979). "Hierarchical and parallel mechanisms in the organization of the visual cortex." *Brain Research Review*, 1, pp. 345-394.

Tieman, S.B., and H.V.B. Hirsch. (1982). "Exposure to lines of only one orientation modifies dendritic morphology of cells in the visual cortex of the cat." *Journal of Comparative Neurophysiology*, 211, pp. 353-362.

Van Essen, D.C., and J.H.R. Maunsell. (1983). "Hierarchical organization and functional streams in the visual cortex." *Trends in Neurosciences*, 6, pp. 1-6.

Vidayasagar, T.R. (1987). "A model of striate response properties based on geniculate anisotropies." *Biological Cybernetics*, 57, pp. 11-23.

Wehner, R. (1981). "Spatial vision in arthropods." In H. Autrum (ed.) *Handbook of Sensory Physiology, VII/ 6C, Vision in Invertebrates*. Berlin: Springer-Verlag, pp. 287-616.

Wilson, H.R., D. Levi, L. Maffei, J. Rovamo, and R. DeValois. (1990). "The perception

of form: Retina to striate cortex." In L. Spillmann and J.S. Werner (eds.) *Visual Perception: The Neurophysiological Foundations.* San Diego: Academic Press, pp. 231-272.

SYMBOLS

Symbol	Description
RX0	retinal sustained off-center element
RX1	retinal sustained on-center element
RY0	retinal transient off-center element
RY1	retinal transient on-center element
LGN	lateral geniculate nucleus of the thalamus
TX0	thalamic sustained principal off-center element
TX1	thalamic sustained principal on-center element
TY0	thalamic transient principal off-center element
TY1	thalamic transient principal on-center element
IX0	thalamic sustained inhibitory off-center element
IX0b	thalamic sustained inhibitory off-center input buffer element
S	superior colliculus saccade signal
4Cα	cortical sustained input element
4Cβ	cortical transient input element
4Bv	cortical vertical simple line detector element
4Bh	cortical horizontal simple line detector element
4Bd$^-$	cortical diagonal (negative slope) simple line detector
4Bd$^+$	cortical diagonal (positive slope) simple line detector
5Ar	cortical right motion detector element
5Al	cortical left motion detector element
5Ad	cortical down motion detector element
5Au	cortical up motion detector element
3Br	cortical right motion complex element
3Bl	cortical left motion complex element
3Bd	cortical down motion complex element
3Bu	cortical up motion complex element
3Bru	cortical diagonal (north-east) motion complex element
3Bld	cortical diagonal (south-west) motion complex element
3Brd	cortical diagonal (south-east) motion complex element
3Blu	cortical diagonal (north-west) motion complex element

REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE February 1993	3. REPORT TYPE AND DATES COVERED Final

4. TITLE AND SUBTITLE
AN ALGORITHM FOR SIMPLE AND COMPLEX FEATURE DETECTION:
from Retina to Primary Visual Cortex

5. FUNDING NUMBERS
PE: 0602936N
WU: DN303020

6. AUTHOR(S)
M. R. Blackburn

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Command, Control and Ocean Surveillance Center (NCCOSC)
RDT&E Division
San Diego, CA 92152-5000

8. PERFORMING ORGANIZATION REPORT NUMBER
TD 2456

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Chief of Naval Research
Arlington, VA 22217

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This report describes a hierarchy of functions for the processing of visual motion and pattern information culminating in descriptive sets of features. Processing elements that model simple feature detectors respond maximally only when a stimulus grating pattern with a specific orientation and spatial period is located in phase with the center and surround of the element's receptive field. Processing elements that model complex feature detectors respond maximally only when a stimulus grating pattern with a specific orientation and spatial period is moved in a specific direction through the element's receptive field.

14. SUBJECT TERMS
center surround receptive fields
log-polar mapping
thalamic LGN
motion selectivity
cortex

15. NUMBER OF PAGES
32

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAME AS REPORT

NSN 7540-01-280-5500

Standard form 298 (FRONT)

21a. NAME OF RESPONSIBLE INDIVIDUAL	21b. TELEPHONE (Include Area Code)	21c. OFFICE SYMBOL
M. R. Blackburn	(619) 553–1904	Code 943

INITIAL DISTRIBUTION